Innovation & Trends
In
Missing Data Sensitivity Analyses

Craig Mallinckrodt (cmallinc@lilly.com)
April 23, 2014

Outline

- **Background**
- **New approaches**
- **Example**

Current State

- NRC panel convened at request of FDA. Strong agreement among stake holders
- Recommendations include
  - Clear objectives
  - Maximize retention
  - Sensible primary analysis supported by sensitivity analyses
  - Need for further research on sensitivity analyses and how to use them to draw inference
    - DIASWG

www.missingdata.org.uk

search... SEARCH

MAIN MENU

- Home
- Introduction to missing data
- FAQs
- Example analyses
- REALCOM Impute and Stata
- Substantive model compatible FCS
- Talks
- **DIA working group**
- Bibliography
- Courses / Workshops
- Web Links
- Contact Us
- Book: MI & its application
- Google Discussion Group

DIA working group

The following pages contain materials available from the Drug Information Association (DIA) working group.

Please note that these pages are still under construction.

- **Descriptive Summaries** (1 Article)
- **Inclusive Modeling Approaches** (4 Articles)
- **Missing Not at Random (MNAR) Methods** (4 Articles)
- **Control-Based Multiple Imputation** (3 Articles)
- **Presentations, Manuscripts and Training Materials** (2 Articles)
- **Example Datasets** (1 Article)

Missing Data Mechanisms

- **MCAR: Conditional on the independent variables in the model, neither observed or unobserved outcomes of the dependent variable explain dropout**
- **MAR: Conditional on the independent variables in the model, observed outcomes of the dependent variable explain dropout, but unobserved outcomes do not**

Missing Data Mechanisms

- **MNAR: Conditional on the independent variables in the model and the observed outcomes of the dependent variable, the unobserved outcomes of the dependent variable explain dropout**
- **Alternatively, in MNAR, conditional on the covariates and observed outcomes, the statistical behavior of the unobserved data is *not* equal to that if it had been observed**
  - **In MAR it is equal**

Missing Data in Clinical Trials

- **Efficacy outcomes are seldom MCAR because the observed outcomes typically influence dropout (discontinue for lack of efficacy)**
- **Trials are designed to observe all the relevant information, which minimizes MNAR data**
- **Hence in the highly controlled scenario of longitudinal confirmatory trials, missing data *may* be mostly MAR – or at least MAR provides a good starting point**
- **MAR often plausible, but never provable. Particular MNAR model never provable either**

General Guidance

- **Strive for validity of MAR**
  - **Maximize retention**
  - **Capture data predictive of drop out**
  - **MAR primary analysis options include**

    MI, likelihood-based, wGEE
- **Assessing consequences of departures from MAR is the cornerstone of sensitivity analyses**

Established MNAR Methods

- **General classes of MNAR methods based on factorizations of the likelihood functions for the joint distribution of the outcome variable and the indicator variable for whether or not a data point is observed**
- **Factorization in this context means that the hypothetical "full" data are split into two parts: the actually observed and the missing**
  - ***Shared parameter model***
  - ***Selection models***
  - ***Pattern mixture models***

Selection Model

- $f(Y_i, R_i|\theta,\Psi) = f(Y_i|\theta)\ f(R_i|Yo_i, Ym_i,\Psi)$

- The joint distribution of the ith patient's outcomes (Y_i) and the missingness indicators (R_i) is factored as the marginal distribution of Y_i and the conditional distribution of R_i given Y_i

- Assume:
  - Full-response data comprise (Y_1,Y_2)
  - Objective is to compare Y_2 means
  - Y_2 is missing on some individuals

Notation

Subject i is to be measured at times j = 1, . . . , n

Y_{ij} = measurement on subject i at time j, j = 1 - n

R_{ij} = indicator variable = 1 if Y_{ij} observed, 0 otherwise.

Group Y_{ij} into a vector $Y_i = (Yo_i, Ym_i)$

Yo_i contains Y_{ij} for which $R_{ij} = 1$,

Ym_i contains Y_{ij} for which $R_{ij} = 0$.

Group R_{ij} into a vector R_i commensurate with Y_{ij} such that all 1s are paired with the Yo_i and the 0s are paired with the Ym_i

Ψ = Parameters describing missingness process

θ = Parameters describing measurement process

Selection Model

- No info on association b/t R and Y_2 because Y_2 is always missing when R = 0
- Can fit model even though no empirical information on some parameters b/c of the parametric and structural assumptions imposed
- Standard recommendations are to input values of this parameter(s) across a plausible range and assess consistency of the difference between group means for Y_2

Caveats

- How do we determine plausible range?
- How do we determine dropout model?
  - Same or different for each treatment?
  - Same or different for each time point
  - MNAR methods are sensitive to model misspecification
- Never have data to inform these choices

Outline

- **Background**
- **New approaches**
- **Example**

Modeling Alternative: Doubly Robust Methods

- **Protects against mis-specified dropout model**
  - **3 Parts: 1) analysis model, 2) dropout model, 3) model for the joint distribution of the partially and fully observed data (compatible with 1)**
  - **If either model 2 or 3 is wrong, but not both, the estimators in 1 are still consistent**
  - **If model 1 is wrong, e.g., because a key confounder is omitted, then estimates of all parameters will typically be inconsistent**
- **Some implementations are complex**
- **Need more experience**
- **Macro available at missingdata.org.uk**

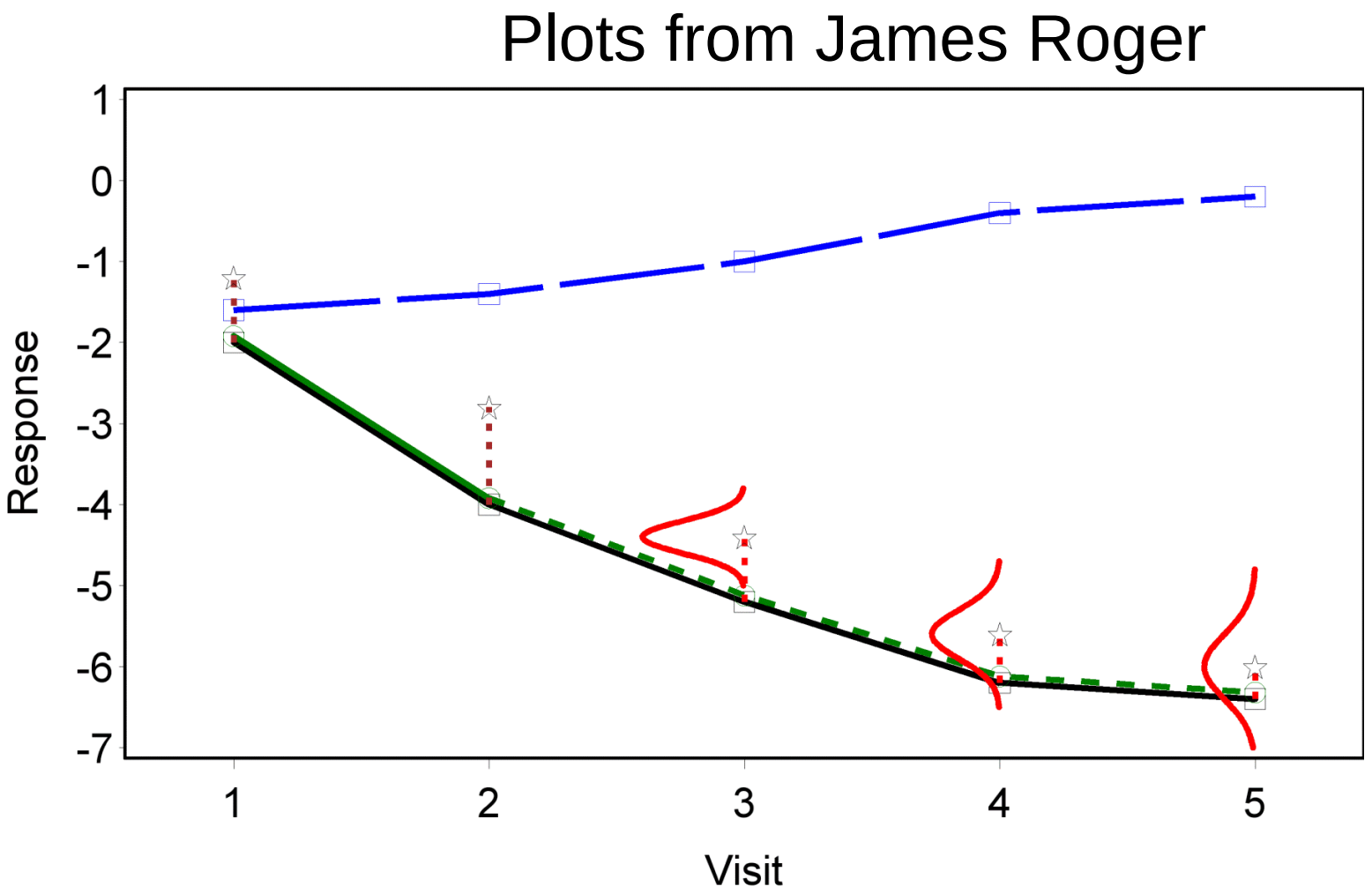
Another Alternative: Controlled Imputations

- **Assumptions can be transparent, debated, and evaluated**
- **Tweak on standard MI, implement w standard software**
- **General idea is to create specific and relevant departures from MAR**
  - **Reference-based: Plausible worst case**
  - **Delta-adjustment: Progressive stress test**
  - **Several implementations of each allows fit for purpose**

Controlled Imputations: Reference Based

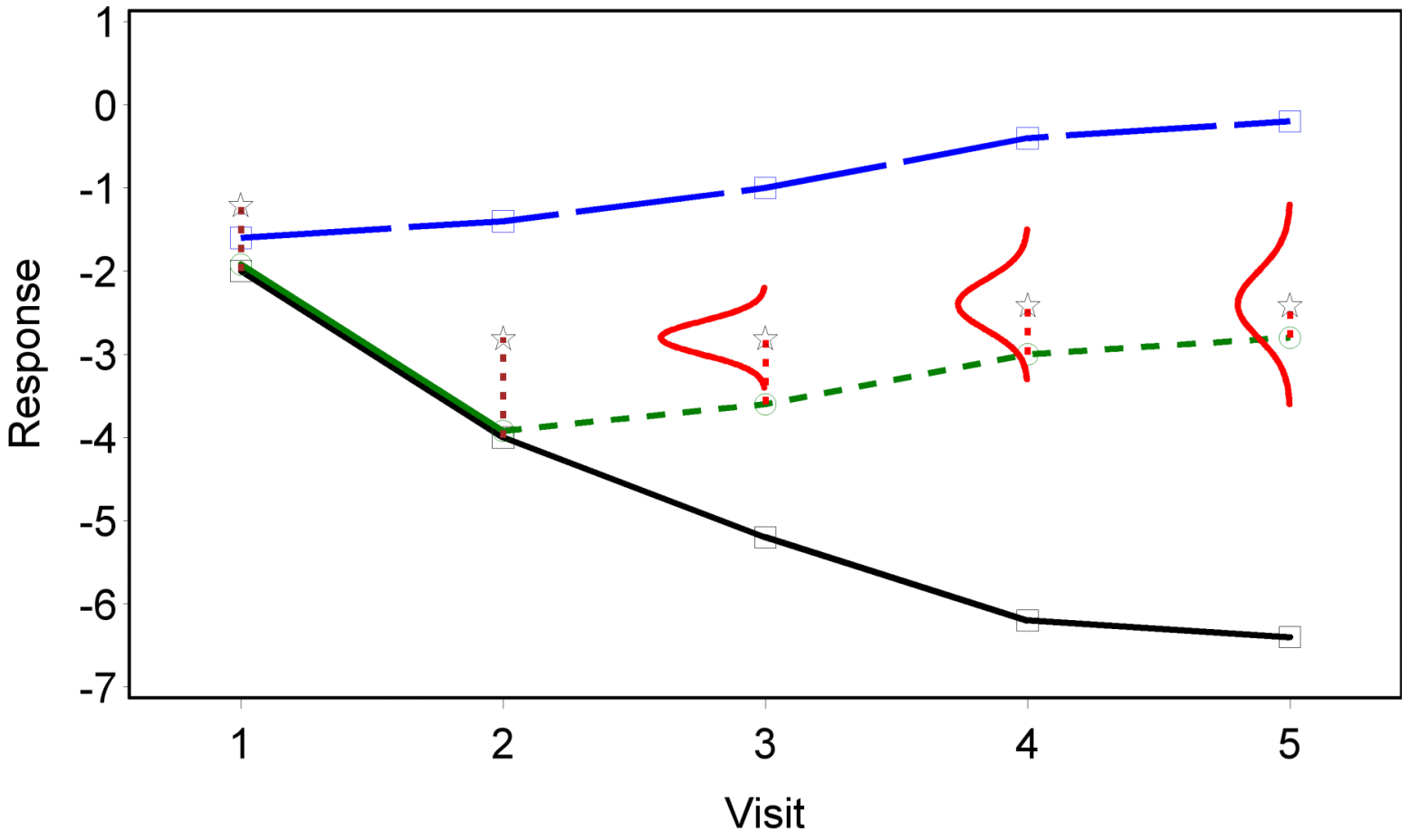
- Copy increment from reference (CIR)
  - After dropout, change for drug = change for placebo
    - Use for disease modifying drugs
- Copy reference (CR)
  - The statistical behavior of drug treated patients after dropout gradually transitions to reference arm
    - Use for drugs with long on target half life
- Jump to reference (J2R)
  - …immediately becomes that of reference patients
    - Use for drugs with short on target half-life

Imputation assuming MAR. Withdrawal after two visits

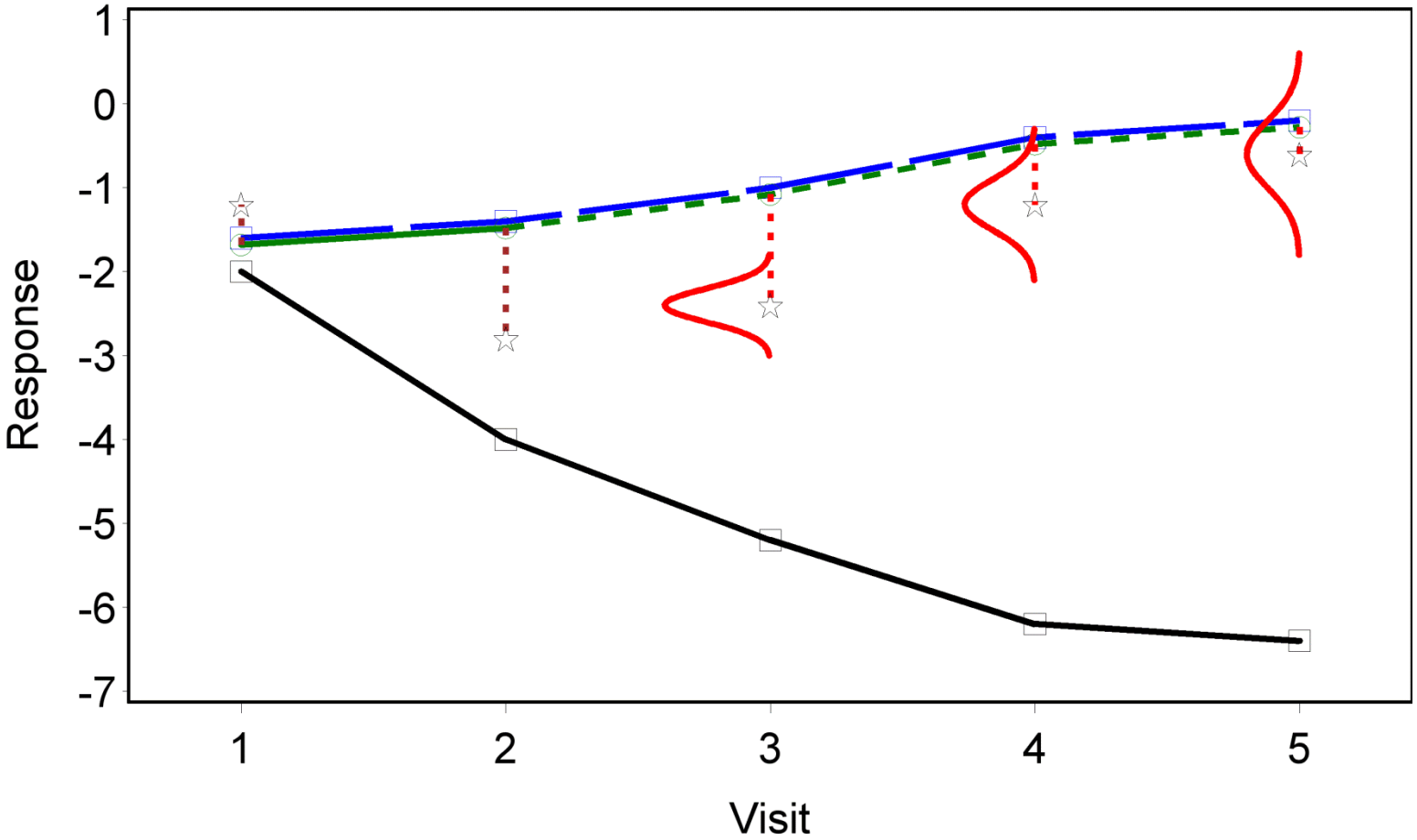


Black line Active. Blue line Reference. Dotted Green line (imputation).

Copy Increment From Reference (CIR)

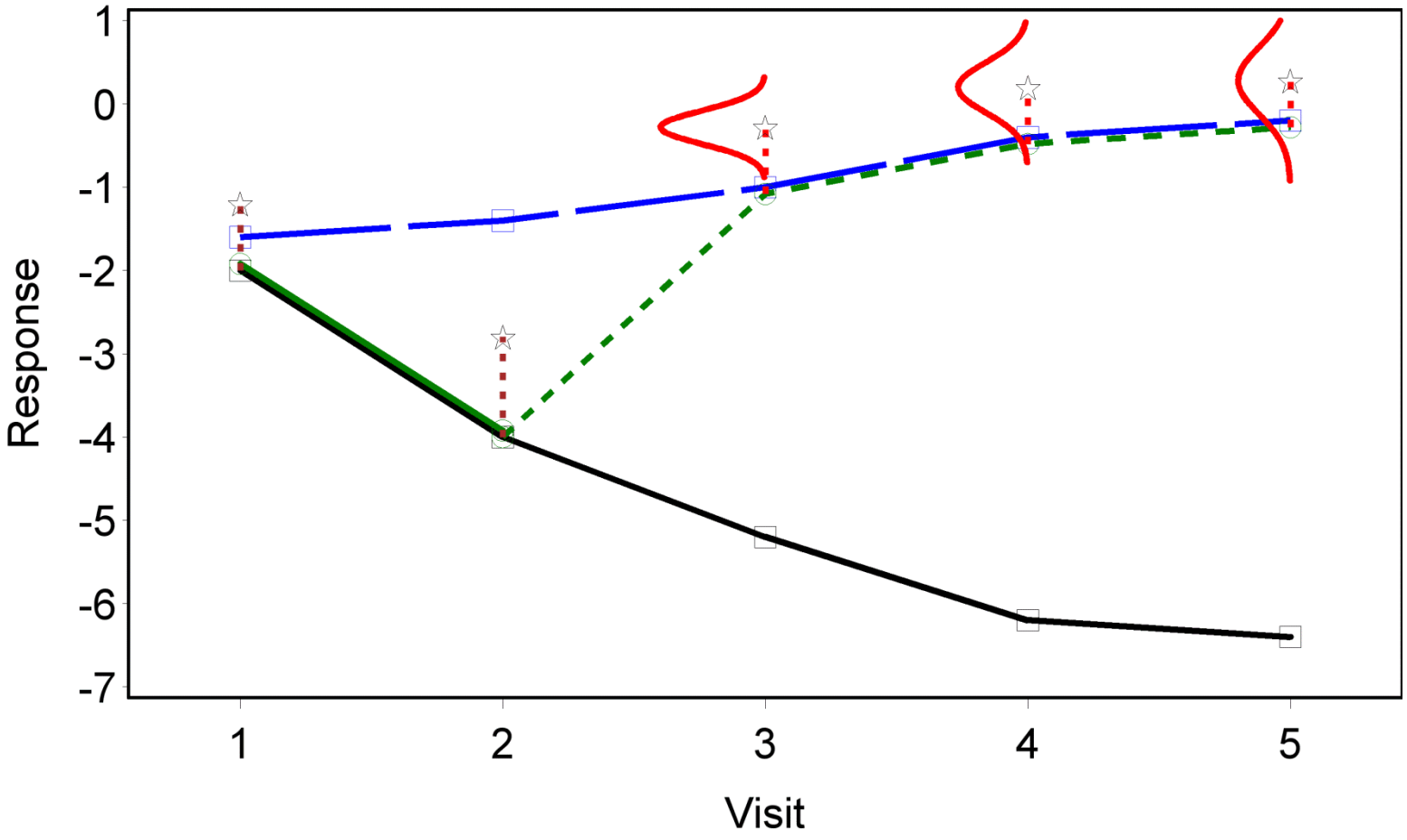


Copy Reference (CR): Residuals prior to withdrawal are based on reference arm mean



Green line shows the means A_{JK} (circles), dotted after withdrawal.
Brown residuals are for two observed values (star) before withdrawal.
Red "residuals" show location of means (star) for conditional distribution.
Red Normal curve indicates actual conditional distribution

Jump to reference (J2R). Residuals based on deviation from assigned treatment arm



Controlled Imputations: Delta Adjustment

- **Subtract a constant (delta) from imputed value at visit X that then further influences imputed values at visit > X**
    - **First missing visit only**
    - **All missing visits**
- **Subtract delta after completion of all imputations**
- **Progressively increase delta until conclusion from primary analysis is overturned**

Outline

- **Background**
- **New approaches**
- **Example**

Example Data from Major Depression

- **Two real but contrived data sets (n=100/arm)**
- **Drug arm patients randomly selected from 3 active arms**
- **Placebo arms mostly as is (with minor replication)**
- **Nearly identical designs**
  - **8-week, double blind, randomized 1:1:1:1**
  - **Assessments @ weeks 1,2,4,6,8**
  - **Similar inclusion / exclusion**
  - **Low dropout from EU study with ext and titration**
  - **High dropout from US fixed dose, no extension**

Example Data

- **High dropout**
  - **Completion rates: placebo 60%, drug 70%**
  - **1000 planned observations, 850 available**
- **Low dropout**
  - **Completion rates: placebo 92%, drug 92%**
  - **1000 planned observations, 961 available**

Therapeutic Innovation and Regulatory Science; 2013, 48(1): 68-80.

Patient Retention

	Week				
	1	2	4	6	8
High dropout					
Placebo	8	7	12	13	60
Drug	9	6	10	5	70
Low Dropout					
Placebo	2	0	3	3	92
Drug	2	1	2	3	92

Prior to week 8 these values are the number of dropouts. For week 8 the values are the number of completers because week 8 is the last scheduled assessment.

Example: Reference-Based Imputation Results

	LSMEANS		LSMEAN	Std	
	Placebo	Drug	Difference[1]	Error	P value
High dropout (35%) dataset[1]					
MAR	-5.95	-8.24	2.29	1.00	0.024
CIR	-5.95	-7.78	1.83	0.97	0.004
CR	-5.96	-7.71	1.75	0.98	0.075
J2R	-5.97	-7.57	1.60	0.99	0.110
Low dropout (8%) dataset[1]					
MAR	-10.56	-12.40	1.84	0.70	0.009
CIR	-10.55	-12.27	1.72	0.70	0.015
CR	-10.55	-12.27	1.72	0.70	0.015
J2R	-10.55	-12.26	1.71	0.70	0.016

Difference MAR vs. J2R over 6x greater in high dropout.

Delta-adjustment Results

	Low Dropout			*High dropout*		
Value of Delta	Endpoint Contrast	Std Error	Pvalue	Endpoint Contrast	Std Error	PValue
0	1.85	0.71	0.009	2.31	1.02	0.024
0.5	1.77	0.71	0.013	2.00	1.03	0.051
2.0	1.52	0.73	0.037			
2.5	1.44	0.74	0.051			

- Delta required to overturn significance of primary result 5x larger for high dropout data set

- Change in endpoint contrast per 1 unit change in Delta
  - Low dropout = 0.16
  - High dropout = 0.62

Example: Specifying Sensitivity Analyses

- **The key assumption made in the primary analysis regarding missing data is that the missingness arises from a missing at random mechanism. Sensitivity of inferences to this assumption will be evaluated by comparing the magnitude of the treatment effect from the primary analysis to…**

Discussion

- **Proper sensitivity analyses allow us to *assess* sensitivity**
- **Lower rates of dropout is the only way to *improve* sensitivity**
- **Controlled imputations are useful and intuitive**
  - **Assumptions can be informed by data obtained from follow-up phase**

Selected Publications

- A structured approach to choosing estimands and estimators in longitudinal clinical trials. (2012). *Pharmaceutical Statistics*; 11:456–461.
- Missing data: Turning guidance into action. (2013). *Statistics in Biopharmaceutical Research*; 5(4): 369-382.
- Recent Developments in the Prevention and Treatment of Missing data. (2013). *Therapeutic Innovation and Regulatory Science*; 48(1): 68-80.
- A structured framework for assessing sensitivity to missing data assumptions in longitudinal clinical trials. (2013). *Pharmaceutical Statistics*; 12:1–6.
- *Preventing and Treating Missing Data in Longitudinal Clinical Trials*. (2013). Cambridge University Press, Cambridge.
- Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models. *Pharmaceutical Statistics*. (2013). DOI: 10.1002/pst.1549
- Analysis of longitudinal trials with missing data: a framework for relevant, accessible assumptions, and inference via multiple imputation. (2013). *Journal of Biopharmaceutical Statistics*, 23: 6, 1352-1371.
- *Clinical Trials with Missing Data: a Guide for Practitioners*. (2014) Wiley, Chichester.
- A Multiple-Imputation-Based Approach to Sensitivity Analyses and Effectiveness Assessments in Longitudinal Clinical Trials. (2014). *Journal of Biopharmaceutical Statistics*, 24: 211–228,

www.missingdata.org.uk

search... SEARCH

MAIN MENU

- Home
- Introduction to missing data
- FAQs
- Example analyses
- REALCOM Impute and Stata
- Substantive model compatible FCS
- Talks
- **DIA working group**
- Bibliography
- Courses / Workshops
- Web Links
- Contact Us
- Book: MI & its application
- Google Discussion Group

DIA working group

The following pages contain materials available from the Drug Information Association (DIA) working group.

Please note that these pages are still under construction.

- **Descriptive Summaries** (1 Article)
- **Inclusive Modeling Approaches** (4 Articles)
- **Missing Not at Random (MNAR) Methods** (4 Articles)
- **Control-Based Multiple Imputation** (3 Articles)
- **Presentations, Manuscripts and Training Materials** (2 Articles)
- **Example Datasets** (1 Article)